MF821 Project: Trading in a Volatile Time

George Kepertis & Julien Bessette

05/01/2020

On March 12, the stocks plunged in their worst day since 1987. Investors became more worried about the rapid spread of the coronavirus around the world and the impacts it would have on different industries. The industries that are the most exposed to this outbreak were hard hit, such as airline and cruise industries. The goal of this report is to put together the techniques we have learned to create trading strategies and discuss which one performed better during the period. More precisely, we will be taking advantage of the extremely volatile market environment we are currently experiencing to design a trading strategy that will, hopefully, be profitable.

1 Data Preparation and Illustration

We choose to use the data starting from January 2, 2020 to April 24, 2020. Our data will capture a period when the markets were calmer and a period when the markets were more volatile. We will take advantage of this when we design our trading strategy. We first calculate the log returns of the S&P 500 index using the adjusted closed prices. The histogram of the log returns is shown in figure 1. We also fitted a line from the normal distribution to illustrate how the log returns are distributed.

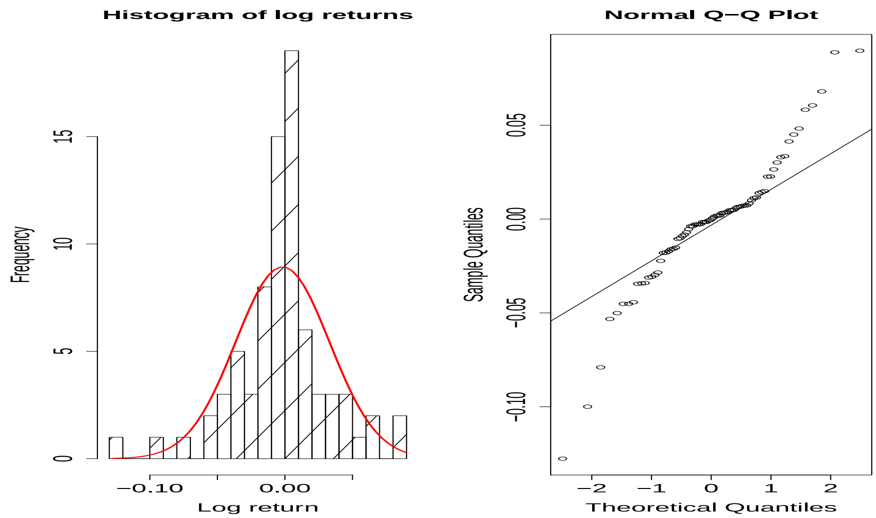


Figure 1: Distribution of the log returns

Lastly, we plot a Normal Q-Q plot of the log returns with the theoretical line as well. From the Q-Q plot, we see that our data has fatter tails than the normal distribution and has a higher kurtosis. This also aligns with what we obtain in the histogram of the log returns.

2 Fitting an Autoregressive Model

2.1 Splitting the data

We split our data into a portfolio formation period and a portfolio testing period. Our portfolio formation period consists of 50 days and the testing period consists of 29 days. Figure 2 shows the plot of the returns over this time period. The red line shows the data we'll use for our portfolio formation period. The rest is part of the testing period. As we can see, most of the data come from the data in January and February, and we also capture data towards the end when the markets began to become more volatile.

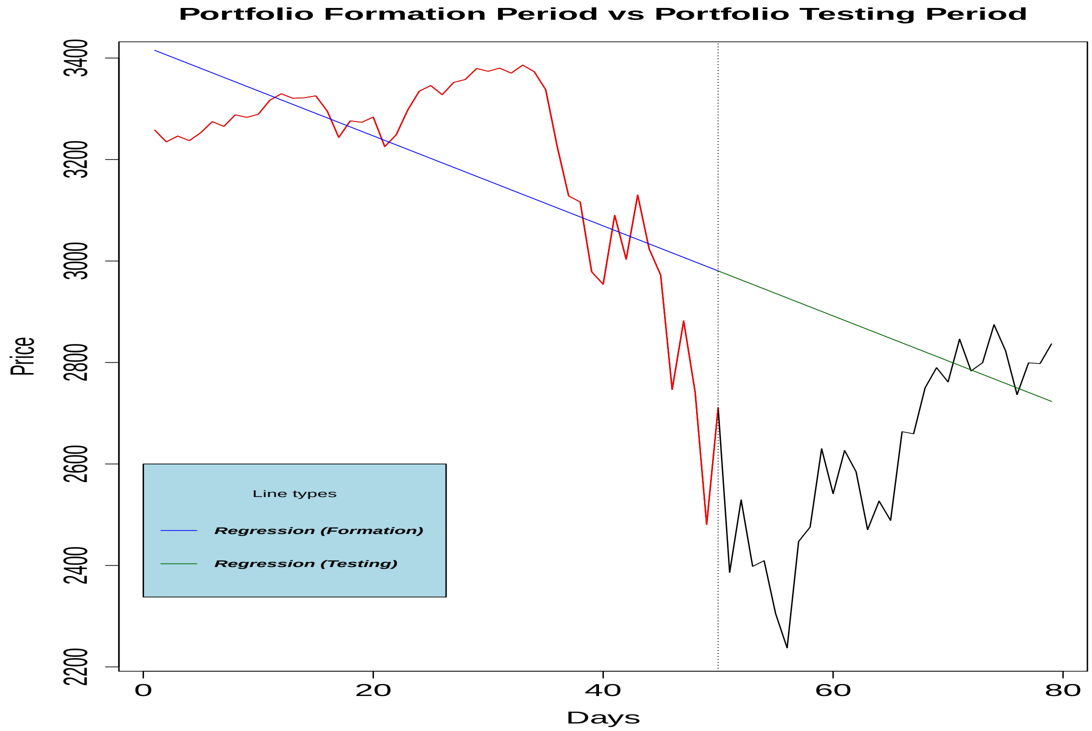


Figure 2: Distribution of the log returns

2.2 Estimating a linear model

We call $S^{\text{ave}}(t)$ our linear regression of the stock prices as a function of time, using the data from the formation period. We use $t = 1, 2, 3 \ldots$, and our regression is written as follows

$$S^{\text{ave}}(t) = 3415.19 - 8.8706t.$$

Having a negative slope coefficient tells us that we do capture the period in the data when the markets start going down. Here, we want our trading strategy to be able to capture the market movements from March. The blue line in figure 2 shows the regression. We also show the cutoff point when it switches to the testing period. The green line is simply the continuation of the regression.

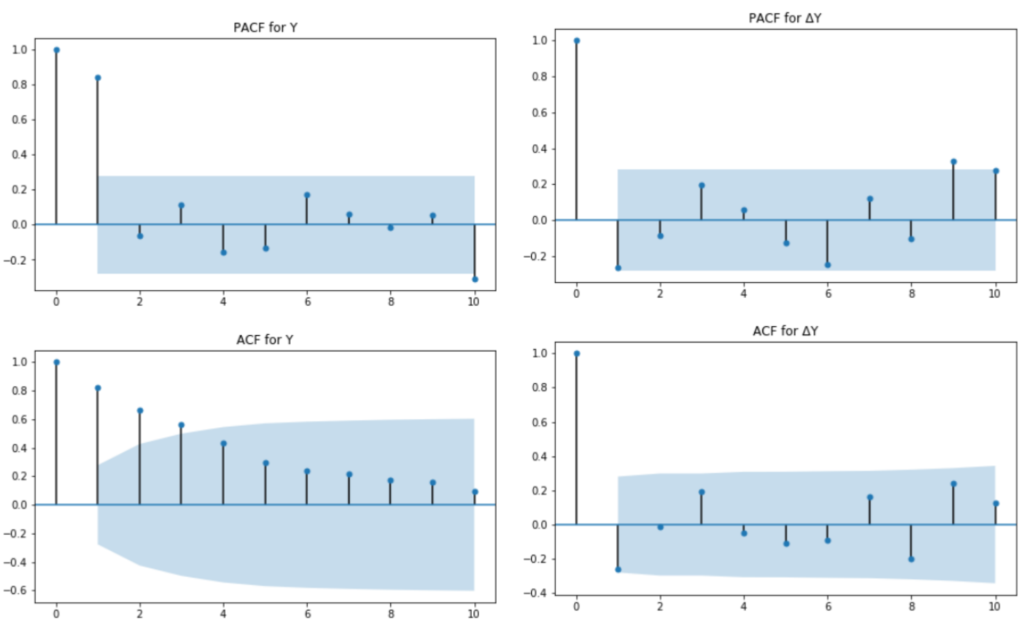
2.3 Fitting AR(1) models

In the portfolio formation, we calculate $Y(t) = S(t) - S^{\text{ave}}(t)$. Here, $Y(t)$ represents price deviations of the S&P 500 from its linear trend.

Since the goal is to extract a continuous time Ornstein-Uhlenbeck process for $Y(t)$, we can look at $AR(1)$ models for $Y(t)$ and $\Delta Y(t)$.To do so, we remove the linear regression fitted values from the real prices of the index, and fit a $AR(1)$ on this data. In order words, to make the times series stationary, we take the number of differencing to be one. Our models are written as

$$Y(t) - Y(t-1) = \mu_1 + \phi_1 (Y(t-1) - Y(t-2)).$$
$$Y(t) = \mu_2 + \phi_2 Y(t-1)$$

We find $\mu_1 = -3.4079651$, $\phi_1 = -0.3091553$, and $\mu_2 = -1.616532$, $\phi_2 = 0.877707$.
Figure 3 shows the ACF and PACF plots for $Y(t)$ and $\Delta Y(t)$. From the figures, we actually see that statonarity analysis on $\Delta Y(t)$ would probably not be very appropriate as there is no significant autocorrelation at short lags. On the other hand, we see that $Y(t)$ does exhibit significant autocorrelation up to and including lag 3. In fact, the most appropriate model for $Y(t)$, would be an $ARMA(1,3)$, but an $AR(1)$ will also capture a lot of its behavior.



(a) ACF and PACF for Y

(b) ACF and PACF for ΔY

Figure 3: Stationarity of $Y(t)$ and $\Delta Y(t)$

2.4 Extracting a continuous time Ornstein Uhlenbeck process

We now extract a continuous time mean reverting model for Y using the $AR(1)$ model from the previous letter. We take a mean revering model of the form

$$\mathrm{d}Y_t = k(\theta - Y_t)\,\mathrm{d}t + \sigma \mathrm{d}W_t. \tag{1}$$

If we discretize (1)

$$\begin{aligned} Y(t) - Y(t-1) &= k(\theta - Y(t-1))\,\Delta t + \sigma(W_t - W_{t-1}) \\ Y(t-1) - Y(t-2) &= k(\theta - Y(t-2))\,\Delta t + \sigma(W_{t-1} - W_{t-2}). \end{aligned}$$

The difference of the two equations above gives

$$Y(t) - Y(t-1) = (1 - k\Delta t)(Y(t-1) - Y(t-2)) + \epsilon_t,$$

where the noise term $\epsilon_t = \sigma(W_t - 2W_{t-1} + W_{t-2})$. From the regression that we had above, it becomes clear that

$$\frac{1-\phi}{\Delta t} = k$$

and, given k,

$$k\theta\Delta t = \mathbb{E}\left[Y(t) - Y(t-1) + kY(t-1)\,\Delta t\right].$$

or equivalently

$$k\theta\Delta t = \mu \Rightarrow \theta = \frac{\mu}{k\Delta t}$$

Using these relations, we find the following:
If the process is estimated based on $\Delta Y(t)$, $k = 1.309155$ and $\theta = 3.751646$.
If the process is estimated based on $Y(t)$, $k = 0.12229$ and $\theta = -13.21847$.
And finally, for the volatility parameter in our model, we use the sample volatility of Y, and we find that $\sigma = 154.05110$.

3 Trading Strategies using ad-hoc bands

In this section, we look at different choices of ad-hoc bands to design the best trading strategy. To design these ad-hoc bands, we assume that the prices will exhibit a mean-reverting behavior. We explore three trading strategies in more detail next.

3.1 Strategy 1

To design our first trading strategy, we use the volatility of $Y(t)$. The idea is to calculate the predictions for the testing period according to the linear model we developed above. This step is pretty simple since the model is linear, the predicted values are the ones that fall on the green line in 2. Using these values, we calculate a rolling mean using a window of three days. The plot below shows the buying and the selling bands, using a window of 3 and a standard deviation of 1 on the left (trading strategy 1 (a)) On the right, we show the same trading strategy, but now we use a rolling window of length 5 and a standard deviation of 0.5 (trading strategy 1(b)). For each day, the two boundaries are given by

$$
\begin{aligned}
Y^{\text{up}}(t) &= S^{\text{ave, rol}}(t) + \text{n.std} * \sigma_Y \\
Y^{\text{down}}(t) &= S^{\text{ave, rol}}(t) + \text{n.std} * \sigma_Y,
\end{aligned}
$$

where $S^{\text{ave, rol}}(t)$ is the rolling mean value of the variables $S^{\text{ave}}(t)$ on that day for the appropriate window length, n.std is the number of standard deviations we're using to design the strategy, and σ_Y is the standard deviation of Y in the testing sample. The graphs are shown in figure 4.

One obvious extension to this model is to have a rolling standard deviation instead of a fixed standard deviation. This would account for the changing volatility as we trade in the testing sample. We will be designing such a strategy next using Bollinger bands. For now let's take a look at what's going on in figure 4. Obviously, our choice of n.std will have an effect on the space between the two bands. Since we capture the dynamics of the prices before the vertical line, we have a downward trend. We see that our two samples are divided in a period of high volatility with a downward price trend, meaning that the model expects the prices to keep going down. Our lines are shifted up because we also take into account the first 20 days of the sample when the market was more stable and was at a higher price level. In our testing sample, shortly after the vertical line, we see that the prices start going back up so that the final prices are around the bands. In other words, the prices do exhibit a pattern of mean reversion which our model predicts. However, one obvious drawback is that we're using a linear regression that does not take into account the new information. This will be accounted for when we use Bollinger bands next. First, let's discuss the actual transactions that take place according to this strategy. Since the price is below the buying price on the first day of the testing period in both graphs, we buy where the first red point is. On the left, since the price does not cross the selling boundary, we close our position at the end of the testing sample. On the right, it crosses where the second red point is, so we sell before reaching the end of the testing period. For trading strategy (a), we end up with a capital of \$ 1121.6. The results trading strategy (b) are shown in 2.

The question is now the following. Can we do better than this very simple and intuitive trading strategy? One obvious choice to consider now is the Bollinger bands. This technical

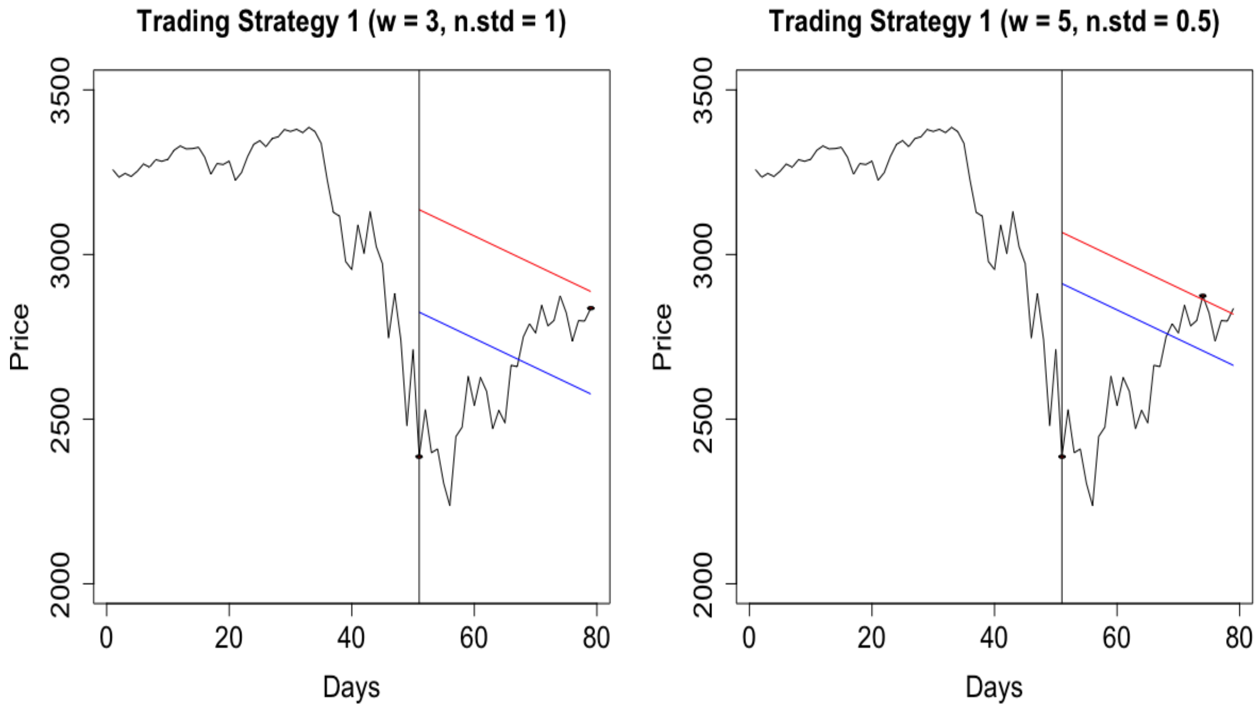


Figure 4: Trading Strategy 1

analysis tool is used by many traders in different markets and can offer a good insight on how the market is behaving. For example, it can be used to determine if the market is overbought or oversold. Here, we will monitor breakouts and trade based on this very simple strategy. But first, let's go over what are Bollinger bands, and what they specifically tells us.

3.2 Strategy 2

The Bollinger bands are in fact three bands: one simple moving average for the middle band, two bands above and below the middle bands. Here, we will calculate these two bands using a moving standard deviation and adding n.std standard deviations to the moving average. For the lower band, we remove n.std standard deviations to the moving average. Here, we use a very simple technique: we buy when the lower band is breached, and we sell when the upper band is breached. We could combine other indicators to our trading strategy, but we consider the very simple case that we are trading only Bollinger bands.

One drawback obviously is that if it's breached again at a lower price, then we will not be able to buy a second time. And so there's the risk that we sell below the buying price.

This assumption obviously could be relaxed (i.e. not invest all of the funds the first time), or other indicators could help signal when it's optimal to trade, but for now let's keep our assumptions in place. The figure with the Bollinger bands is shown below.

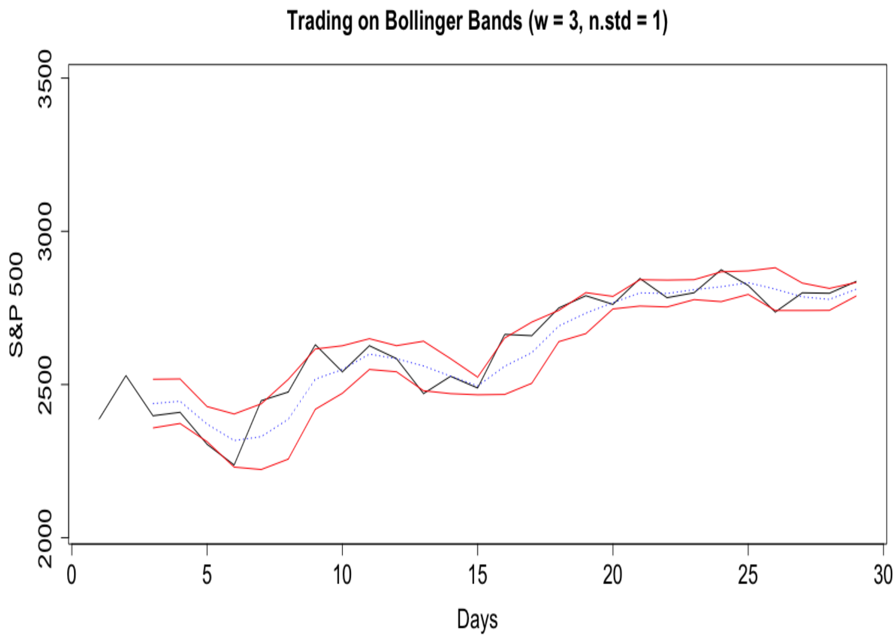


Figure 5: Trading on Bollinger bands

Table 1 summarizes the transactions over the testing period based on this strategy. If

Price	**Transaction**
2304.92	Buy
2447.33	Sell
2470.50	Buy
2663.68	Sell
2736.56	Buy
2836.74	Sell

Table 1: Transactions based on the Bollinger bands strategy

we assume that our starting capital is \$1000 and we make the transactions above by buying fractions of the asset, we end up with a capital of \$ 1186.72.

Let's compare this strategy do the buy and hold strategy. We buy at the beginning of the sample, when the price was \$ 2386.13, 0.419 units of the S&P. We sell it back at the end of the testing period, when the price is at \$ 2836.84, to end up with a capital of \$ 1188.846. The buy and hold strategy performs slightly better than trading on Bollinger bands.

3.3 Strategy 3

We can also design a trading strategy based on the $AR(1)$ model that we built in section 2. Again, our regression was written as

$$Y(t) - Y(t-1) = \mu + \phi(Y(t-1) - Y(t-2)).$$

Here, we keep the $AR(1)$ model constant, we don't reestimate the parameters as we include more days. An obvious extension to this (we could do it?) is to reestimate the model as we go, but here, to be consistent with the model we use above, we will keep parameters constant. So using this regression we can predict $\Delta(t+1)$, using this prediction, we can predict $\Delta(t+2)$, and so on. For every days in the testing sample, we take the actual value of Y that happened on that day, and use the AR model to predict the Y for the next day. To obtain the data for the price predictions, we use a rolling mean with a window of 3 days to calculate S^{ave}. With these predictions for the prices, we can calculate the mean and the standard deviation. Finally, we can draw a lower and an upper band. The plot of the actual returns in the testing sample is shown in the next plot. The buying band is shown in blue, and the selling band is shown in red.

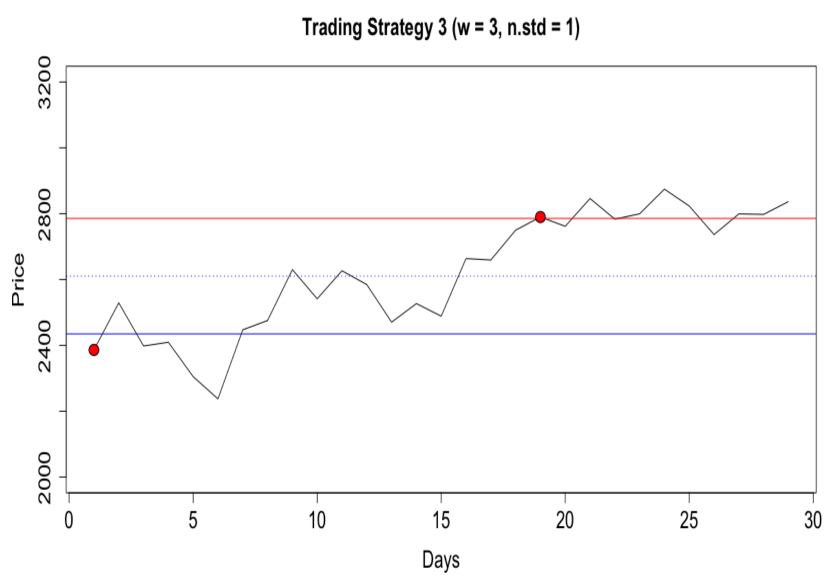


Figure 6: Trading Strategy 3

The plot shows that we buy once. We buy at a price of \$ 2386.13 and close our position at a price of \$ 2789.82, where we see the second red point. With an initial capital of \$ 1000, we end up with \$ 1169.182. The precise results are shown in figure 2.

Some points are worth noting here. It is obvious from the plot that it is hard to do better than the buy and hold strategy here, since the price keeps moving upwards during the whole period. So buying at the beginning of the period and closing the position at the end is very optimal in our context.

4 Trading Strategies using Optimal Bands

4.1 Dynamics of the process and the HJB equation

Since we have extracted the dynamics of $Y(t)$ as

$$dY = k(\theta - Y)dt + \sigma dW$$

we will proceed with directly finding the boundary in terms of $Y(t)$. The optimal selling problem is

$$H(y,t) = \sup_{\tau \leq T} \mathbb{E}_{s,t}\left[S_{\tau\wedge T} - c\right] = \sup_{\tau \leq T} \mathbb{E}_{s,t}\left[Y_{\tau\wedge T} + S^{\text{ave}}_{\tau\wedge T} - c\right]$$

The HJB equation is given by

$$\partial H(y,t) + \sup_{\tau \leq T}\left(\left(\mathcal{A}H\right)(y,t)\right) = 0.$$

And we know that $H(y,t)$ satisfies

$$0 = \max\left(Y_t - S^{\text{ave}}_t - c - H(y,t), \partial H(y,t) + \left(\mathcal{A}H\right)(y,t)\right)$$

with terminal condition $H(y,T) = Y_T + S^{\text{ave}}_T - c$. We have the following free boundary equations. In the continuation domain $(Y < Y^*)$,

$$\begin{gathered} \partial H(y,t) + \mathcal{A}_t H(y,t) = 0 \\ Y_t + S^{\text{ave}}_t - c - H(y,t) < 0. \end{gathered}$$

In the termination domain $Y < Y^*$, we have

$$Y_t + S^{\text{ave}}_t - c - H(y,t) = 0,$$

with the smooth fit condition $H'(y,t) = 1$, for $y = y^*$.

We now turn to the optimal buying problem. We have

$$G(y,t) = \sup_{\eta \le T} \mathbb{E}_{s,t}\left[H\left(y_{\eta\wedge T}, \eta \wedge T\right) - Y_{\eta\wedge T} - S^{\text{ave}}_{\eta\wedge T} - c\right].$$

The HJB equation is given by

$$0 = \partial G(y,t) + \sup_{\eta \le T}\left(\left(\mathcal{A}_t G\right)(y,t)\right).$$

We have that $G(y,t)$ satisfies

$$0 = \max\left(H(y,t) - y - S^{\text{ave}}_t - c - G(s,t), \partial G(y,t) + \left(\mathcal{A}_t G\right)(y,t)\right),$$

with terminal condition $G(y,T) = H(s,T) - y - S^{\text{ave}}_T - c = -2c$. We have the following free boundary equations. In the continuation domain ($Y > \hat{Y}$),

$$\begin{aligned} \partial G(y,t) + \mathcal{A}_t G(y,t) &= 0 \\ H(y,t) - y - S^{\text{ave}}_t - c - G(y,t) &< 0. \end{aligned}$$

In the termination domain ($Y < \hat{Y}$),

$$\begin{aligned} \partial G(y,t) + \mathcal{A}_t G(y,t) &\le 0 \\ H(y,t) - y - S^{\text{ave}}_t - c - G(y,t) &= 0 \end{aligned}$$

with the smooth fit condition $G'(y,t) = 1$ for $Y = \hat{Y}$.
Discretizing:

$$\begin{aligned} &\partial H(y,t) + \left(\mathcal{A}_t H\right)(y,t) \\ &\Rightarrow \partial H(y,t) + k(\theta - y)\,\partial_y H(y,t) + \frac{\sigma^2}{2}\partial_y^2 H(y,t) = 0 \end{aligned}$$

Using the Euler–Maruyama discretization method, we get

$$\begin{aligned} \frac{H(y_i,t_j) - H(y_i,t_{j-1})}{h_t} = k(\theta - y_i)\,\frac{H(y_{i-1},t_j) - H(y_{i-1},t_j)}{2h_y} \\ + \frac{\sigma^2}{2}\frac{H(y_{i+1},t_j) - 2H(y_i,t_j) + H(y_{i-1},t_j)}{h_y^2}. \end{aligned}$$

The goal is to isolate $H(y_{i-1}, t_j)$ in the expression above. We have

$$H(y_i, t_{j-1}) = H(y_{i-1}, t_j) \underbrace{\left(\frac{\sigma^2 h_t}{2h_y^2} - \frac{k(\theta - y_i) h_t}{2h_y} \right)}_{l_i} + H(y_i, t_j) \underbrace{\left(1 - \frac{\sigma^2 h_t}{h_y^2} \right)}_{a_i} + H(y_{i+1}, t_j) \underbrace{\left(\frac{k(\theta - y_i) h_t}{2h_y} + \frac{\sigma^2 h_t}{2h_y^2} \right)}_{u_i}.$$

We use the explicit scheme by going back in the grid to find at which price trading is optimal for every time t.
Using the Euler–Maruyama discretization method for the process G yields the equation

$$G(y_i, t_{j-1}) = G(y_{i-1}, t_j) \underbrace{\left(\frac{\sigma^2 h_t}{2h_y^2} - \frac{k(\theta - y_i) h_t}{2h_y} \right)}_{l_i} + G(y_i, t_j) \underbrace{\left(1 - \frac{\sigma^2 h_t}{h_y^2} \right)}_{a_i} + G(y_{i+1}, t_j) \underbrace{\left(\frac{k(\theta - y_i) h_t}{2h_y} + \frac{\sigma^2 h_t}{2h_y^2} \right)}_{u_i}.$$

And again, we use the explicit scheme by going back in the grid to find at which price trading is optimal, for every t. We need to use the same grid than with H since the continuation and the termination domains will depend on the value $H(s_i, t_j)$ at each node.

4.2 Computing the optimal boundary

First, we need to choose a grid that satisfies the condition $|a_i| < 1, \forall i < M$, where M is the number of subdivisions of the variable Y in our grid. This will ensure that the update matrix in the PDE scheme will not "explode" if we take powers of it.
We can do this using the following grid:

$$\begin{aligned}
&T = 29, \text{the number of testing days} \\
&N = 40000, \text{the number of subdivisions of time} \\
&y_{max} = -y_{min} = 3000 \\
&M = 1000, \text{the number of subdivisions of Y}
\end{aligned}$$

which implies

$$h_t = \frac{T}{N} = 0.000725$$
$$h_y = \frac{2Y_{max}}{M} = 6$$

so $|a_i| = 0.522069 < 1$ which is acceptable for our scheme.
With a specific grid, we are now ready to solve the PDE and find the optimal boundary. More details regarding the actual calculation, including how the update matrix is defined, can be found in the code.

4.3 Trading Strategy

The scheme gives us an optimal boundary for the variable $Y(t)$. We also calculate the variable S_t^{ave} in a rolling window of 3 days and use those values to get back values of the price $S(t)$. Then, if the price is below the lower boundary $Y_l(t) + (S^{\text{ave}}(t))_{roll}$, we buy the stock, and when the price exceeds the upper boundary $Y_d(t) + (S^{\text{ave}}(t))_{roll}$, we sell the stock. Our results can be summarized in figure 7 below:

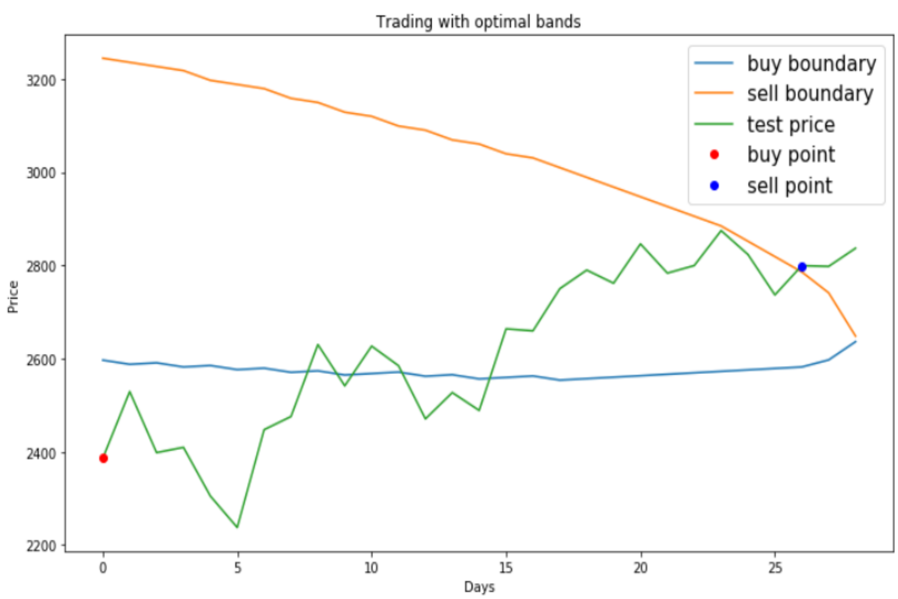


Figure 7: Trading with optimal bands

When we do the trading on the same day, our profit is \$ 173.16, when we delay the transactions by a day the profit is \$106.20, and finally if we short sell to finance the strategy we lose \$ 15.69.

5 Summary of all trading strategies

We show in panel A of table 2 the details of the trading strategies we developed above. In panel B, we look into more details how a delay in the order by one day will have an impact on the performance of our portfolio. We see that that the returns decrease when we delay the trading by one day, as we observe in the last rows of the two panels. The performances of trading strategy 1 (b) and trading strategy 3 are particularly affected by the delay of execution. Trading strategy 2 is not affected as much, which could be explained by the fact that the trading is split into 3 buy/sell transactions. Also, there is not a big influence on the performance of the optimal band strategy.

We also calculate the annualized Sharpe ratio of our P&L. Denote by $\mathbf{PL} = (PL_t)_{t\in\mathcal{J}}$, where $\mathcal{J} = \{1, 2, \ldots, 29\}$ is the indexes of the days in our testing sample, and PL_t is the profit/loss on that day. We define the annualized Sharpe ratio as

$$\text{Sharpe Ratio}_{\text{ann}} = \sqrt{252} \times \frac{\mathbb{E}\left[\mathbf{PL}\right]}{\sqrt{\mathbb{V}\left[\mathbf{PL}\right]}}.$$

We see that the values of the Sharpe ratios decrease when we have delay in the trading.

Results of the trading strategy				
	Trading Strategy 1 (b)	Trading Strategy 2	Trading Strategy 3	Optimal Bands
Panel A: No Delay				
Ending Capital	$ 1204.7	$ 1186.72	$ 1169.182	$ 1173.16
Number of transactions	2	6	2	2
Mean (P & L)	7.310555	6.668588	6.042214	6.18425
Sharpe Ratio$_{\text{ann}}$	3.168754	4.435461	2.720095	2.62666
Returns of trading period	0.2046956	0.1867205	0.169182	0.117316
Panel B: With Delay				
Ending Capital	$ 1116.231	$ 1180.037	$ 1091.903	$ 1106.20
Number of transactions	2	6	2	2
Mean (P & L)	4.151103	6.429893	3.282247	3.79300
Sharpe Ratio$_{\text{ann}}$	1.967189	4.44682	1.635148	1.77652
Returns of trading period	0.1162309	0.180037	0.09190292	0.110620

Table 2: Results of the trading strategies, with and without delay

To our surprise, when there is no delay, the very first trading strategy that we developed, which was based on the linear regression, performs the best amongst the four, followed by the Bollinger bands strategy. When we introduce delay in trading, we observe that the Bollinger bands strategy performs the best. We saw before that if we followed the buy and hold strategy over the testing period, we made a return of 11.88 %. The only strategy that beats the buy and hold is the trading strategy 2, as shown in table 2.

6 Conclusion

Downturns in markets happen frequently, and setbacks are usually followed by recovery, but investors should still know how to be profitable during these more volatile times and develop reliable techniques to trade when markets experience more extreme disturbances. In our short paper, we analyzed three trading strategies based on ad-hoc bands to develop reliable trading strategies. Then we looked at the underlying stochastic process of the asset prices to develop a trading strategy based on optimal bands.

A possible extension to our work would be to assume different dynamics for the underlying asset prices and develop a new trading strategy based on these new dynamics. Also, instead of fitting a linear regression, we could fit a quadratic regression to the asset prices.

Statement of contribution

This is a statement confirming that both Julien Bessette and George Kepertis worked equally on the project, including writing code, and typing up results.

Instructions for running the code attached and where to find results

Using the ˆGSPC3.csv file, the code for questions 1,2,3 can be found in `R`. The code for question 4 can be found in a `Python` Jupyter Notebook. It is important that everything is run in the order that it comes and several of the results are printed when running individual cells.